An iterative multilevel method for computing wavefields in frequency-domain seismic inversion

Yogi A. Erlangga* and Felix J. Herrmann

SUMMARY

We describe an iterative multilevel method for solving linear systems representing forward modeling and back propagation of wavefields in frequency-domain seismic inversions. The workhorse of the method is the so-called multilevel Krylov method, applied to a multigrid-preconditioned linear system, and is called multilevel Krylov-multigrid (MKMG) method. Numerical experiments are presented for 2D Marmousi synthetic model for a range of frequencies. The convergence of the method is fast, and depends only mildly on frequency. The method can be considered as the first viable alternative to LU factorization, which is practically prohibitive for 3D seismic inversions.

INTRODUCTION

Imaging of the earth's subsurface can computationally be done in the frequency domain by inverting seismic data using the adjoint-state method; see, e.g., Plessix (2006). One advantage, among others, of doing inversion in the frequency domain is the possibility of inversion with only a small subset of random frequencies, as shown, e.g., in Sirgue and Pratt (2004) and Mulder and Plessix (2004). In relation to this, it is demonstrated by Lin et al. (2008) (in this proceedings) that it is possible to recover the wavefield with only 30 per cent of frequencies required by the classical Nyquist sampling theory, by solving a sparsity-promoting recovery problem in, e.g., the Fourier domain. In combination with less number of grid points required to resolve the wavefield, seismic inversion in the frequency domain will now require less complexity as compared to the time domain migration.

Translated to the seismic language, the adjoint-state method consists of forward modeling, in which a wave equation is solved for a given velocity background and source position, and computing a correction to the given velocity background. The correction is determined by minimizing the misfit between the seismic data, recorded at the receiver positions, and the forward model data at the positions corresponding to the physical receivers. The gradient of the misfit functional is closely related to the migrated image in geophysical applications, and can be computed by multiplication of forward and back propagated wavefields. These wavefields are computed via a finite-difference approximation of the wave equation.

As mentioned, e.g., in Pratt et al. (1998), wavefield computations may be done by direct methods, by first constructing the LU factors of the finite-difference wave equation matrix $\mathbf{A}$. Once these factors are available, for a given angular frequency ω and velocity background $c(x)$, both forward modeling and back propagation problems can be solved using the same LU factors for a set of shot-receiver positions.

Computing the LU factor of $\mathbf{A}$ is, however, very costly and memory consuming. In two dimensions, for example, computing LU factors require $O(n_x^2 n_z)$ work and $\sim 2n_x^2 n_z$ storage. In three dimensions, work and storage needed by an LU factorization becomes $O(n_x^3 n_y^3 n_z^3)$ and $2n_x^2 n_y^2 n_z$, respectively. Considering these facts, it is considered prohibitively expensive to do frequency-domain seismic inversion based on LU factorization.

Iterative methods are usually considered as an alternative to direct methods (Saad (2003)). An iterative method relies mostly on matrix-vector multiplications with $\mathbf{A}$. Hence, an iterative method is less demanding in terms of memory and computational complexity. Iterative methods are, however, generally less robust as compared to direct methods, and in particular for wave simulations, all efficient, *ad hoc* iterative methods fail to converge for frequency as high as 5 Hz, in the case of, e.g., the 2D Marmousi synthetic model. Erlangga et al. (2006) proposed an iterative method based on Krylov subspace method, preconditioned by a damped acoustic wave operator. The preconditioner is inverted approximately by one multigrid iteration. In this paper, we call this method the MG method. While the method improves the convergence significantly, the convergence still shows dependence on angular frequency.

In this paper, we will show that the convergence can be further improved if the MG method is used within the context of the multilevel Krylov method, called the MK method, recently introduced in Erlangga and Nabben (2007). We call this combination "the MKMG method".

THEORY

For frequency-domain seismic imaging, the wavefield can be modeled by the frequency-domain acoustic wave equation

$$\mathcal{H}u(\omega, x_s; x) := -\left(\nabla \cdot \nabla - \frac{\omega^2}{c(x)^2}\right) u(\omega, x_s; x) = b, \quad (1)$$

equipped with radiation boundary conditions, where $\omega = 2\pi f$, with f the frequency in Hz, and shot position x_s. Application of finite differences on Equation 1 results in the linear system

$$\mathbf{A}[\mathbf{c}]^s \mathbf{u}^s = \mathbf{b}^s, \quad \mathbf{A} \in \mathbb{C}^{n \times n}, \quad (2)$$

where the vector $\mathbf{u}$ is a discrete approximate solution to the solution function u evaluated on the finite-difference grid points. In seismic inversions, we are concerned with the (real) velocity background associated with the wavefield data recorded at the receiver positions. Given a reasonable initial velocity background $\mathbf{c}$ and the corresponding solution vector $\mathbf{u}$ of the forward model 2, the real velocity background can be estimated by minimizing the functional $F = \frac{1}{2}\langle \delta \mathbf{d}, \delta \mathbf{d} \rangle$, where $\delta \mathbf{d} := \mathbf{u}(\omega, x_s; x_r) - \mathbf{d}(\omega, x_s; x_r)$ defines the misfit between the forward model wavefield at the receiver positions, $\mathbf{u}(\omega, x_s; x_r)$, and the seismic data $\mathbf{d}(\omega, x_s; x_r)$. This minimization amounts to an update

$$\mathbf{c}^{\ell+1} = \mathbf{c}^\ell - \alpha \nabla_{\mathbf{c}} F, \quad (3)$$

where $\nabla_c F = real\{\mathbf{J}^H \delta \mathbf{d}\}$, and $\mathbf{J} = \mathbf{J}(\mathbf{u}, \mathbf{c})$ the Jacobian matrix of $\mathbf{u}$ with respect to $\mathbf{c}$. Computing the Jacobian matrix is very costly, and in practice, only the action of the Jacobian on $\delta \mathbf{d}$ is

Multilevel method in seismic inversion

needed. This is equivalent to back-propagating waves from the receiver positions. In this case, the back-propagated wavefield $\mathbf{v}^s$ is computed from

$$(\mathbf{A}^s)^H \mathbf{v}^s = \delta \mathbf{d}^s, \qquad (4)$$

with $(\mathbf{A}^s)^H$ the adjoint of $\mathbf{A}^s$ in Equation 2. For a set of shots $s = 1, \ldots, S$, the gradient can be written as

$$\nabla_{\mathbf{c}} F = \frac{\omega^2}{c^3} diag(\mathbf{U}^H \mathbf{V}), \qquad (5)$$

where $\mathbf{U}^H = [\mathbf{u}^1 \ldots \mathbf{u}^S]$ and $\mathbf{V}^H = [\mathbf{v}^1 \ldots \mathbf{v}^S]$. The term $\nabla_{\mathbf{c}} F$ therefore corresponds to the migrated image in seismic imaging.

By virtue of Equation 5, a frequency-domain seismic imaging now amounts to computing forward and back-propagated wavefields from Equations 2 and 4, respectively. In the next section, we describe an iterative method used to perform these computations.

Methods

The main driver for our method is the Krylov subspace method. It is well known that the convergence of a Krylov method depends to some extend on the eigenspectrum of the given matrix. In particular, if the condition number of $\mathbf{A}$, denoted by $\kappa(\mathbf{A})$, is small (or more specific, close to one), then the convergence of a Krylov method can be expected to be fast. For Equations 2 and 4, $\kappa(\mathbf{A})$ is typically very large, called ill-conditioned. In order to have a better conditioning, an equivalent, right preconditioned system is solved, namely

$$\mathbf{A}\mathbf{M}^{-1}\hat{\mathbf{u}} = \mathbf{b}, \qquad \mathbf{u} = \mathbf{M}^{-1}\hat{\mathbf{u}}, \qquad (6)$$

where $\mathbf{M}$ is a preconditioning matrix. Note that we have dropped the superscript s to simplify the notation.

For some class of matrices, the matrix $\mathbf{M}^{-1}$ is usually formed as a sort of approximation to $\mathbf{A}^{-1}$. But, following Erlangga et al. (2006), $\mathbf{M}$ can be based on a finite-difference approximation of the damped, or shifted, wave operator

$$\mathscr{M} := -\nabla \cdot \nabla - \frac{\omega^2}{c(x)^2}(1 - \beta i), \quad i = \sqrt{-1}, \qquad (7)$$

with $\beta > 0$. In this case, the matrix $\mathbf{M}$ is not an approximation of $\mathbf{A}$ in the usual sense. But, considering the preconditioned Equation 6, the matrix $\mathbf{A}\mathbf{M}^{-1}$ has the following spectral properties. First, the eigenvalues of $\mathbf{A}\mathbf{M}^{-1}$ are enclosed by a circle with center along the real axis of the complex plane. The center and the radius of the circle depends on the choice of β, but this circle never touches lines $real = 0$ and $real = 1$. This is sketched in Figure 1 (b). Hence, compared to the eigenspectrum of $\mathbf{A}$, shown in Figure 1 (a), the eigenspectrum of $\mathbf{A}\mathbf{M}^{-1}$ is very much clustered. Furthermore, while the real part of eigenvalues of $\mathbf{A}$ changes sign (called indefinite), the real part of eigenvalues of $\mathbf{A}\mathbf{M}^{-1}$ all have the same sign. For a given β, the enclosing circle is independent of the finite-difference grid size h and angular frequency ω. Secondly, the largest eigenvalue of $\mathbf{A}\mathbf{M}^{-1}$ in magnitude is bounded above by one. Thirdly, the smallest eigenvalue lies very close to zero at a

distance $O(\omega^{-1})$. The first two spectral properties guarantee h-independent convergence, and by the last property we may expect convergence mildly dependent of ω.
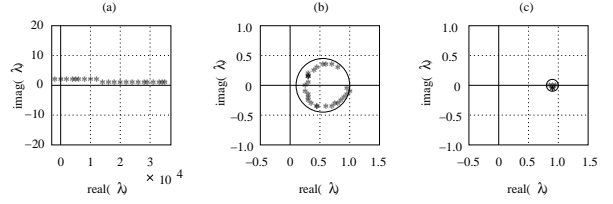

Figure 1: Illustration of the spectrum of (a) $\mathbf{A}$, (b) $\mathbf{A}\mathbf{M}^{-1}$, and (c) $\mathbf{A}\mathbf{M}^{-1}\mathbf{Q}$. The asterisks " " show the eigenvalues.

It is worth noting that in practice $\mathbf{M}^{-1}$ is never computed exactly. Rather, in preconditioning steps in a Krylov subspace method, the action of $\mathbf{M}^{-1}$ on a vector is done approximately by one multigrid iteration. In this case, the preconditioning steps can be carried out with only $O(n_x n_z)$ work in 2D and $O(n_x n_y n_z)$ work in 3D. Details of this technique are discussed in Erlangga et al. (2006) and Riyanti et al. (2006a), where Bi-CGSTAB (van der Vorst (1992)) is used for the Krylov iteration. Throughout the paper, Bi-CGSTAB preconditioned by multigrid will be in short referred to as the MG method.

As the largest eigenvalue is always bounded above by one, it appears that the convergence of the MG method is now solely determined by the small eigenvalues of $\mathbf{A}\mathbf{M}^{-1}$, which is of order ω^{-1}. Hence, convergence can be further improved if these small eigenvalues are somehow shifted to some values far from zero. This shift should however be performed such that the upper bound of the eigenspectrum remains the same. Such a "spectral modification" is done by the so-called multilevel Krylov (MK) method, introduced recently by Erlangga and Nabben (2008) and Erlangga and Nabben (2007). Define a projection $\mathbf{Q}$ as

$$\mathbf{Q} := \mathbf{I} - \mathbf{Z}\mathbf{E}^{-1}\mathbf{Y}^T \mathbf{A}\mathbf{M}^{-1} + \lambda_n \mathbf{Z}\mathbf{E}^{-1}\mathbf{Y}^T, \qquad (8)$$

with $\mathbf{E} := \mathbf{Y}^T \mathbf{A}\mathbf{M}^{-1}\mathbf{Z} \in \mathbb{C}^{m \times m}$, λ_n the largest eigenvalue of $\mathbf{A}\mathbf{M}^{-1}$ in magnitude, and $\mathbf{Z}, \mathbf{Y} \in \mathbb{R}^{n \times m}$, $m \ll n$, any rank m projection matrices. In this case, $n = n_x n_z$ in 2D, or $n = n_x n_y n_z$ in 3D. A convergence acceleration of a Krylov method is obtained if it is applied to the system

$$\mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\tilde{\mathbf{u}} = \mathbf{b}, \quad \mathbf{u} = \mathbf{M}^{-1}\mathbf{Q}\tilde{\mathbf{u}}. \qquad (9)$$

The effect of the inclusion of $\mathbf{Q}$ can be briefly explained as follows. Denote the eigenspectrum of $\mathbf{A}\mathbf{M}^{-1}$ by

$$\sigma(\mathbf{A}\mathbf{M}^{-1}) = \{\lambda_1, \ldots, \lambda_n\} \in \mathbb{C}, \qquad (10)$$

where λ's are ordered such that $|\lambda_1| \leq |\lambda_2| \cdots \leq |\lambda_n|$. In this case, $|\lambda_n| \to 1$, and $|\lambda_1| = O(\omega^{-1})$. For any full rank $\mathbf{Z}, \mathbf{Y}$, the eigenspectrum of $\mathbf{A}\mathbf{M}^{-1}\mathbf{Q}$ becomes

$$\sigma(\mathbf{A}\mathbf{M}^{-1}\mathbf{Q}) = \{\mu_1, \ldots, \mu_n\} \in \mathbb{C}. \qquad (11)$$

Because $|\lambda_n| \to 1$, in this case $|\mu_n| \to 1$ and $|\mu_i| = O(\lambda_n) = O(1)$. This is illustrated in Figure 1 (c), which shows that the eigenvalues of $\mathbf{A}\mathbf{M}^{-1}$ are now moved towards 1, due to the

inclusion $\mathbf{Q}$. Since μ_i now is $O(1)$, the condition number of $\mathbf{AM}^{-1}\mathbf{Q}$ can be expected to be $O(1)$. In terms of number of iterations, this will lead to a much faster convergence. The application of a Krylov method on Equation 9 results in the multilevel Krylov-multigrid method, or MKMG.

A brief discussion on the implementation aspects of the MKMG method is given in Appendix A. For more details, see Erlangga and Nabben (2008) and Erlangga and Nabben (2007).

EXAMPLE AND DISCUSSION

Numerical experiments are performed based on a part of 2D Marmousi synthetic model as shown in Figure 2. We consider two situations encountered in frequency-domain inversions: forward and back-propagation model, governed by Equations 2 and 4, respectively. To mimic forward modeling, a point source is located at the position $(3000, 50)$ m. For back propagation, we consider point sources distributed along the receiver line at $z = 50$ m. In both cases, $f = [5, 30]$ Hz is used.
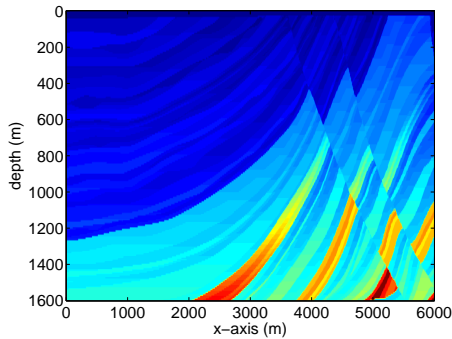

Figure 2: Part of 2D Marmousi data. $c_{\min} = 1500$ m/s, $c_{\max} = 4450$ m/s.

As aforementioned, in the MG method, Bi-CGSTAB is applied on Equation 6. For MKMG, we apply GMRES (Saad and Schultz (1986)) on Equation 9. In this case, both methods require two matrix-vector multiplications with $\mathbf{A}$ and two preconditioner solves by one multigrid iteration per iteration; thus, they amount to approximately the same complexity. We set $\mathbf{Y} = \mathbf{Z}$, with $\mathbf{Z}$ associated with the bilinear interpolation of the coarse grid into the fine grid. The iteration is terminated if the initial residual has been reduced by six orders of magnitude. A snapshot of forward modeling at $f = 20$ Hz is shown in Figure 3. We note that it is theoretically difficult to determine the optimal β in Equation 7 for MKMG. In this case, we have to rely on numerical experiments. We found that $\beta = 1$ is the best value we could obtain so far. This value of β is different from the MG method, where $\beta = 0.5$.

Convergence results of MKMG for forward and back-propagation problems are shown in Figure 4 and Table 1, respectively, with the grid size $h \approx \frac{1}{18} c_{\min}/f$. The results indicate that there is practically no significant difference in convergence for forward and back-propagation problems. With $\beta = 1$, the convergence of MKMG is practically independent of frequency.
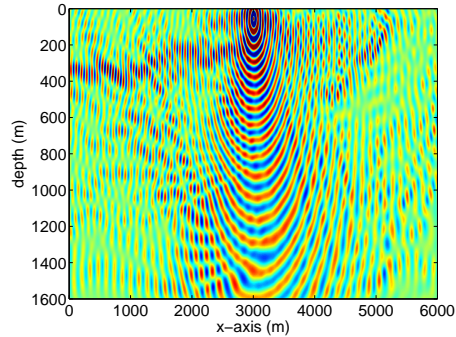

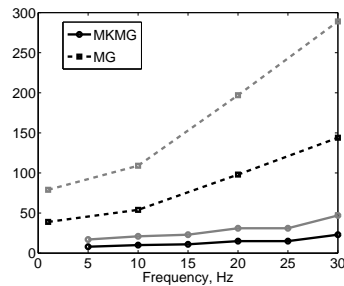Figure 3: Wavefield from forward modeling. $f = 20$ Hz.

Figure 4: Forward modeling using 2D Marmousi velocity model. Black lines indicate the number of iterations; the gray lines are the number of multiplications with $\mathbf{A}$.

In Figure 4, we also compare the convergence results of MKMG with the MG method. In terms of the number of iterations, MKMG converges about five times faster than the MG method. It is worth mentioning that without any preconditioner, Bi-CGSTAB applied to Equation 2 with $f = 1$ Hz requires 17446 iterations to converge.

Table 2 compares the convergence results of MKMG in Figure 4 with the case where the grid size is fixed. In this case, for $f = [5, 30]$ Hz we set $h = \frac{1}{18}(c_{\min}/30)$. Hence, for low frequencies the grid used is extremely fine. The table clearly shows that the convergence is independent of h.

Frequency (Hz)	5	10	15	20	25	30
Iteration count	8	10	11	15	15	23

Table 1: Number of MKMG iterations for back propagation using 2D Marmousi velocity model.

Finally, we estimate the total memory used in MKMG and compare it with direct methods in the case of *in-core* implementation. (MKMG can however be implemented without any matrices stored in memory.) For MKMG, we need to store the matrices $\mathbf{A}$'s, $\mathbf{M}$'s, and $\mathbf{Z}$'s (see Appendix A). In addition, since we base the MKMG method on GMRES, we also need to store vectors of length $n_x n_z$ used in GMRES. As seen in Figure 5, MKMG requires less memory than LU factorization.

Frequency (Hz)	5	10	15	20	25	30
Grid adapted to f	8	10	11	15	15	23
Grid fixed at $f = 30$ Hz	8	8	11	12	15	23

Table 2: Number of MKMG iterations for forward modeling using the 2D Marmousi data.

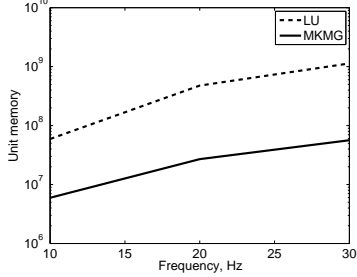This advantage will become more significant in three dimensions.

Figure 5: Memory used in LU factorizations and MKMG.

CONCLUDING REMARKS

Parallel implementations of the MG method have been presented for 2D and 3D forward modeling in Kononov et al. (2006) and Riyanti et al. (2006b), respectively. Since Krylov methods are well parallelizable, this implies that the MKMG method is also well parallelizable.

Riyanti et al. (2006b) note that the 2D frequency-domain migration based on direct methods can be about one order of magnitude faster than the 2D time-domain migration. For direct methods with nested dissection, computing LU factors and linear system solves amount to $n_f O(n^3) = n_f n O(n^2)$ and $n_f n_s O(n^2 \log n)$ work, respectively, with n_f the number of frequencies, n_s the number of shots, and $n = n_x = n_z$. Thus, if $n_s = n$, the complexity is $W_D = n_f n_s(O(n^2) + O(n^2 \log n))$. In direct methods, however, the equality constant for the first term is usually large. For MKMG, the complexity is $W_{MKMG} = n_f n_s n_{it} O(n^2)$, with n_{it} the number of MKMG iterations. Since for $n > e$, $n^2 \log n > n^2$, we see that $W_{MKMG} < C n_f n_s n^2 \log n$ if the corresponding equality constant times n_{it} is smaller than $C \log n$. Considering its fast convergence, it is very likely that for 2D problems, MKMG is more efficient than the most efficient direct methods. This efficiency of MKMG will be more pronounced in 3D. Combining MKMG with sparsity-promoting recovery, which allows the use of a very limited number of frequencies, will further widen the margin between the frequency-domain and time-domain migration.

ACKNOWLEDGMENTS

APPENDIX A

THE MKMG METHOD

To describe the MKMG method, we rewrite Equation 9 as

$$\mathbf{A}_1 \mathbf{M}_1^{-1} \mathbf{Q}_1 \tilde{\mathbf{u}} = \mathbf{b}, \quad \mathbf{u} = \mathbf{M}_1^{-1} \mathbf{Q}_1 \tilde{\mathbf{u}}, \qquad \text{(A-1)}$$

with $\mathbf{Q}_1 := \mathbf{I}_1 - \mathbf{Z}_{12}\mathbf{E}_2^{-1}\mathbf{Y}_{12}^T\mathbf{A}_1\mathbf{M}_1^{-1} + \lambda_n \mathbf{Z}_{12}\mathbf{E}_2^{-1}\mathbf{Y}_{12}^T$, and $\mathbf{E}_2 := \mathbf{Y}_{12}^T\mathbf{A}_1\mathbf{M}_1^{-1}\mathbf{Z}_{12}$. Note that the matrix $\mathbf{Q}_1$ must not be constructed explicitly because $\mathbf{Q}_1$ is dense. The reason of choosing of subscripts in this manner will be clear during the discussion. Suppose that we have $\mathbf{Y}_{12}$ and $\mathbf{Z}_{12}$.

In a Krylov method, the solution subspace is expanded by computing the new basis $\mathbf{w}^{i+1}$ using the previous basis $\mathbf{w}^i$, i.e.,

$$\mathbf{w}^{i+1} = \mathbf{A}_1 \mathbf{M}_1^{-1} \mathbf{Q}_1 \mathbf{w}^i = \mathbf{A}_1 \mathbf{M}_1^{-1} \left(\mathbf{w}^i - \mathbf{Z}_{12}\mathbf{E}_2^{-1}\mathbf{Y}_{12}^T \hat{\mathbf{w}}^i \right), \text{(A-2)}$$

where $\hat{\mathbf{w}}^i = (\mathbf{A}_1 \mathbf{M}_1^{-1} - \lambda_n \mathbf{I}_1)\mathbf{w}^i$. Define $\hat{\mathbf{y}} := \mathbf{Y}_{12}^T\hat{\mathbf{w}}^i$, with $\hat{\mathbf{y}} \in \mathbb{C}^m$. We then need to compute $\hat{\mathbf{z}} = \mathbf{E}_2^{-1}\hat{\mathbf{y}}$. We avoid an explicit computation of $\mathbf{E}^{-1}$ because we need to form $\mathbf{M}_1^{-1}$, which, in our case, is not available because we only know the action of it on a vector via one multigrid iteration. Rather, we will only consider the action of $\mathbf{E}^{-1}$ on $\hat{\mathbf{y}}$ via the implicit equation

$$\mathbf{E}_2 \hat{\mathbf{z}} = \hat{\mathbf{y}}. \qquad \text{(A-3)}$$

To do this, we introduce an approximation to $\mathbf{E}_2$ of the form

$$\mathbf{E}_2 \approx \mathbf{Y}_{12}^T\mathbf{A}_1\mathbf{Z}_{12}(\mathbf{Y}_{12}^T\mathbf{M}_1\mathbf{Z}_{12})^{-1}\mathbf{Y}_{12}^T\mathbf{Z}_{12}. \qquad \text{(A-4)}$$

The products $\mathbf{Y}_{12}^T\mathbf{A}_1\mathbf{Z}_{12} =: \mathbf{A}_2$, $\mathbf{Y}_{12}^T\mathbf{M}_1\mathbf{Z}_{12} =: \mathbf{M}_2$ and $\mathbf{Y}_{12}^T\mathbf{Z}_{12} =: \mathbf{B}_2$ are the Galerkin approximation associated with the individual matrices $\mathbf{A}_1$, $\mathbf{M}_1$, and $\mathbf{B}_1 = \mathbf{I}_1$, respectively, and have a very similar form with the coarse-grid matrix in multigrid. Hence, the matrices $\mathbf{Y}_{12}$ and $\mathbf{Z}_{12}$ can be built in the same way the multigrid interpolation and restriction matrix. Since $m = dim\mathbf{A}_2 = dim\mathbf{M}_2 = dim\mathbf{B}_2 \ll dim\mathbf{A}_1 = n$, we say that the vectors $\hat{\mathbf{w}}$, and respectively $\hat{\mathbf{z}}$, belong to a higher and lower dimension subspace. All matrices which map a vector into the largest subspace are denoted by the subscript "1", etc. To get an accurate $\hat{\mathbf{z}}$ as fast as possible, we can introduce an operator $\mathbf{Q}_2$, which is similar to $\mathbf{Q}_1$. By virtue of approximation A-4, $\hat{\mathbf{z}}$ is solved from Equation A-3 by solving the system

$$\mathbf{A}_2 \mathbf{M}_2^{-1} \mathbf{B}_2 \mathbf{Q}_2 \tilde{\mathbf{z}} = \hat{\mathbf{y}}, \quad \hat{\mathbf{z}} = \mathbf{Q}_2 \tilde{\mathbf{z}}. \qquad \text{(A-5)}$$

By an application of a Krylov method on Equation A-5, a recursive MKMG method results. An MKMG method with four levels is illustrated in Figure A-1.
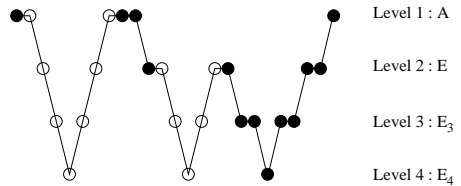

Figure A-1: One MKMG iteration. Black circles indicate the MK step, the white circles indicate the MG loop.

Multilevel method in seismic inversion

REFERENCES

Erlangga, Y. A. and R. Nabben, 2007, On multilevel projection krylov method for the preconditioned helmholtz system: Electronic Transaction in Numerical Analysis, submitted.

———, 2008, Multilevel projection-based nested Krylov iteration for boundary value problems: SIAM Journal on Scientific Computing, to appear.

Erlangga, Y. A., C. W. Oosterlee, and C. Vuik, 2006, A novel multigrid-based preconditioner for the heterogeneous Helmholtz equation: SIAM Journal on Scientific Computing, **27**, 1471–1492.

Kononov, A. V., C. D. Riyanti, S. W. de Leeuw, C. Vuik, and C. W. Oosterlee, 2006, Numerical performance of parallel solution of heterogeneous 2d Helmholtz equation: Computing and Visualization in Science, online http://www.springerlink.com/content/6q377635t271624m/.

Lin, T. T. Y., E. Lebed, Y. A. Erlangga, and F. J. Herrmann, 2008, Interpolating solutions of the Helmholtz equation with compressed sensing: Presented at the SEG Meeting, this proceedings.

Mulder, W. A. and R. E. Plessix, 2004, How to choose a subset of frequencies in frequency-domain finite-difference migration: Geophysical Journal International, **158**, 801–812.

Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applciations: Geophysical Journal International, **167**, 495–503.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Riyanti, C. D., Y. A. Erlangga, R. E. Plessix, W. A. Mulder, C. Vuik, and C. W. Oosterlee, 2006a, A new iterative solver for the time-harmonic wave equation applied to seismic problems: Geophysics, **71**, E57–E63.

Riyanti, C. D., A. Kononov, Y. A. Erlangga, C. Vuik, C. W. Oosterlee, R. E. Plessix, and W. A. Mulder, 2006b, A parallel multigrid-based preconditioner for the 3d heterogeneous high-frequency Helmholtz equation: Journal of Computational Physics, **224 (1)**, 431–448.

Saad, Y., 2003, Iterative methods for sparse linear systems: SIAM.

Saad, Y. and M. H. Schultz, 1986, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems: SIAM Journal on Scientific Computing, **7(12)**, 856–869.

Sirgue, L. and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

van der Vorst, H. A., 1992, Bi-CGSTAB: a fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems: SIAM Journal on Scientific Computing, **13(2)**, 631–644.